Journal of Soft Computing in Civil Engineering 1-1 (2017) 29-53

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

No-Deposition Sediment Transport in Sewers Using Gene Expression Programming

I. Ebtehaj[1], **H. Bonakdari**[2*]

1. Ph.D. Candidate, Department of Civil Engineering, Razi University, Kermanshah, Iran
2. Professor, Department of Civil Engineering, Razi University, Kermanshah, Iran

Corresponding author: *bonakdari@yahoo.com*

https://doi.org/10.22115/SCCE.2017.46845

ARTICLE INFO

Article history:
Received: 26 May 2017
Revised: 13 June 2017
Accepted: 14 June 2017

Keywords:
Bed load;
Sediment transport;
Sewer;
No-deposition;
Gene-expression programming.

ABSTRACT

The deposition of the flow of suspended particles has always been a problematic case in the process of flow transmission through sewers. Deposition of suspended materials decreases transmitting capacity. Therefore, it is necessary to have a method capable of precisely evaluating the flow velocity in order to prevent deposition. In this paper, using Gene-Expression Programming, a model is presented which properly predicts sediment transport in the sewer. In order to present Gene-Expression Programming model, firstly parameters which are effective on velocity are surveyed and considering every of them, six different models are presented. Among the presented models the best is being selected. The results show that using verification criteria, the presented model presents the results as Root Mean Squared Error, *RMSE*=0.12 and Mean Average Percentage Error, *MAPE*=2.56 for train and *RMSE*=0.14 and *MAPE*=2.82 for verification. Also, the model presented in this study was compared with the other existing sediment transport equations which were obtained using nonlinear regression analysis.

How to cite this article: Ebtehaj I, Bonakdari H. No-deposition sediment transport in sewers using gene expression programming. J Soft Comput Civ Eng 2017;1(1):29–53. https://doi.org/10.22115/scce.2017.46845.

1. Introduction

Transmitting flow through sewage channel is often accompanied by solid materials. Sediment deposition takes place because of solid materials wide range entry into the sewer as well as the intermittent and variable nature of flow regimes within the sewer. Therefore, the management of sediment transport in sewers is considered one of the most important items in sewer designing and operation. During wet weather flow, the flow rate is enough to suspend the solid sediments. Solid materials deposition in sewers takes place especially in low flow rate cases such as the beginning of the layout period, low consumption hours or warm seasons of the year. Permanent deposition on the pipe bed causes cross-sectional variation and bed roughness and therefore velocity, and shear stress distribution change and sewer hydraulic resistance consequently influences sediment transport capacity and finally causes operation maintenance cost increments. In order to convey minimum entry flow into the sewer, the slope ought to be as much to be able to prevent sediment deposition, or for a fixed channel slope, the minimum transmitting flow rate shall be as much to be able to transport solid materials. In addition, to design the velocity that is somehow capable of transmitting no-deposition solid materials, pipe diameter shall be selected in a way that transmitting maximum flow rate becomes possible.

Therefore, methods are needed to manage deposit transmission in a way that the transmitting flow would be capable of cleansing deposited sediments. Also, hand design process needs to be economical and optimized [1]. The traditional method of designing sewage channels to prevent sediment deposition in the flow uses minimum velocity or minimum shear stress. In this method, sewer designing was done by presenting a fixed velocity or minimum shear stress at a determined flow depth or specified period. For example, ASCE [2] proposes the constant velocity for full and semi-full flow equal to 0.6 m/s for sanitary sewer and 0.9 m/s for storm sewer. British Standard [3] proposes 0.75 m/s for full flow and storm sewer and 1 m/s for combined sewer. European Standard [4] considers the constant velocity for pipes with diameters less than 300 mm equal to 0.7 m/s. While this criterion has not presented any suggestion for larger diameter pipes, flow conditions are not denoted in this standard. Also, for constant shear stress criterion, ASCE [2] has proposed shear stress within the range of 1.3 to 12.6 $N/m^{2,}$ and Lysne [5] has proposed shear stress between 2 to 4 N/m^2. Therefore, we can conclude that velocity or minimum shear stress values are not equal in different conditions and countries. This is related to implemented experiments, size of sediments in a different region and other parameters. So, in order to determine self-cleansing velocity, one has to achieve factors effective on sediment transport such as sediment concentration and size, flow hydraulic depth or radius, pipe roughness, and diameter, so that the designer can achieve minimum required velocity according to regional conditions.

To survey sediment transmission in the no-deposition case within sewers, sediment transport has been presented in two general ways: using dimensional analysis and semi-experimental relations. In order to present sediment transport relations with the use of dimensional analysis, dimensionless parameters are determined after implementing various experiments and studying the effect of effective parameters on sediment transport, and finally, sediment transport relations are presented. In order to present semi-experimental relations with the use of effective forces exerted on a particle in an equilibrium state, relations are presented. Using dimensional analysis, presented relations are given in three different states. The first approach evaluates densimetric Froude Number (Fr) with the use of volumetric sediment concentration (C_V), relative flow depth (d/R) and overall sediment friction factor (λ_s) [6–9]. The second approach calculates Fr similar to the first case, but the difference is that in this approach in addition to the presented parameter in the first case, dimensionless particle number (D_{gr}) is also used [10–13]. The third approach of the presented relations which uses dimensional analysis evaluates Fr by using volumetric sediment concentration (C_V) and flow proportional depth (d/R or d/y) [11,14,15]. Semi-experimental relations are also presented in different ways and will be briefly presented. May [16] obtained his model of bed load transport based on effective loads which are exerted on particles transmitted at the limit of deposition. Using dimensional analysis, the author simplified the theoretical model in order to present his model and fit it with experimental data. May et al. [17] modified the relation of May [16] by using seven different sets of data. This relation is considered as the best sediment transport relation at the limit of deposition, which is achieved semi-experimentally [18]. Correcting the relation by Ackers and White [19] and in order to consider flow cross section form in pipes, Ackers [20] presented his relation. May [17] presented his relation in a semi-experimental way to transport at the limit of deposition based on effective shear stress on the sediments surface. To develop a new practical methodology for sewer, a comprehensive research project conducted in the UK based on available experimental knowledge. The results of this project were offered by Butler et al. [21]. The harvest of this study is presented as a self-cleansing sewer design methodology based on a new definition of self-cleansing. The authors considered an efficient self-cleansing sewer which has sediment transport capacity by considering a minimum amount of deposited bed to balance between consolidated expenses of construction, operation, and maintenance. Banasiak [22] investigated the behavior of non-cohesive and partly cohesive deposited sediment in a partially full sewer pipes and its effect on the hydraulic performance of sewer. They found the presence of cohesive- like beds is more desirable than granular ones in terms of the bed roughness. Because the attendance of fine sediments as deposited sediment results in partly cohesive deposited solids so that decrease or in such cases prevent bed forms development. Ota and Perrusquia [23] conducted several experimental tests in two sewer pipes at the limit of deposition condition to the measurement of sediment particle and sphere velocity. As regards, sediment transport depends on sediment repose angle; the authors developed a new semi-theoretical equation based on a reclaimed non-dimensional bed shear stress. Safari et al. [24] carried out a series of experiment tests on trapezoidal channel cross-section. Using these samples and collected a wide range of experimental data of U-shape, rectangular and circular channel cross sections from the literature, the authors developed a self-cleansing model based on the definition of a shape factor to consider the effect of channel cross section.

In recent years, using soft computing (SC) in different sciences has led to desirable results [25–29]. To overcome the uncertainty and complexity accompanied with bed load sediment transport estimation in sewers, Azamathulla et al. [13] presented multi-nonlinear regression-based model and adaptive neuro-fuzzy inference systems (ANFIS). They found that the offered ANFIS model could employ as a strength alternative tool in sediment transport prediction at the clean pipe. Ebtehaj and Bonakdari [30] evaluated the performance of artificial neural network (ANN) in the estimation of sediment transport using self-cleaning concept. They found the superior results of ANN in compared with existing regression-based methods. Ebtehaj and Bonakdari [31] employed two different algorithms; back-propagation (BP) and hybrid of back-propagation and least-square (BP-LS); to train ANFIS in predicting of sediment transport in sewers. Moreover, to the generation of fuzzy inference systems (FIS), sub-clustering (SC) and grid partitioning (GP) were utilized. Based on these methods, they introduced four different methods of ANFIS training. The results illustrated that a combination of GP and Hybrid results in the most precise sediment transport prediction.

All computational methods have different advantages and disadvantages depending on the type of problems, the decision on whether or not to use it. In ANN, the learning and computations are easy, but the major drawbacks of this approach are as arriving at the local minimum, less generalizing performance, over-fitting problem and slow convergence speed. Moreover, attaining the optimal structure of a constructed ANN is not simple [32]. The main shortcoming of fuzzy logic (FL) is in finding the shape of each variable, and suitable membership functions are untangled by trial and error [33,34]. To overcome the disadvantage of ANN and fuzzy logic, ANFIS has been introduced which are known as a most popular strong SC tool. ANFIS is an adaptive fuzzy system which allows to the utilization of ANN topology with FL simultaneously.

It not only contains the features of both approaches but also removes some shortcomings of their lonely-utilized case. Indeed, ANFIS consists of ANN advantages such as understanding mathematical details is not obligate and acquaintance with the job data is enough, employed different algorithm within learning course and solving nonlinear complex problems with strong capacity [35]. Moreover, the advantages of ANFIS in comparison with ANN have attained highly nonlinear mapping, better learning capacity, and involves fewer tunable parameters. However, the most constraints in ANFIS are more complex than FIS, not exist for all types of FIS [32] and there is no law for tuning the membership functions.

In addition to these drawbacks, the main problem in both of ANN and ANFIS is the existence of a black-box and don't provide a certain equation to apply in practical applications. Therefore, it needs to a technique to overcome this shortcoming. One of the newest presented models in soft computing topic is Gene Expression Programming (GEP). The main shortcomings of this method are premature convergence due to the derivation of this method from genetic programming and genetic algorithm, preservation of best individual based on roulette-wheel selection method with elitism so that results in losing other better individuals [36] and CPU time-consuming. Azamathulla and Ab. Ghani [37] predicted pipeline scoured depth with the use of GEP and concluded that in comparison with existing models, the presented model provides better results. Khan et al. [38] used GEP to predict bridge pier scour. The authors compared their presented model with artificial neural network and regression relations and concluded that the presented model leads to more satisfactory results when compared to existing models. Chang et al. [39] compared three different methods available in soft computing, adaptive neuro-fuzzy inference system, feed-forward neural network, and GEP, to survey bed load in the rivers. Azamathulla and Ahmad [40] used GEP model to predict transverse mixing coefficient in open channels flow. Using laboratory results mostly, the authors presented a relation to estimate transverse mixing coefficient which presented the results with more precession compared with the existing relations. With the use of Gene-Expression Programming (GEP) in this study, sediment transport in sewerage channels has been studied. The presented model is applicable to the no-deposition case.

To increase the accuracy of the presented model in this study – in comparison with the existing models [41] which only used the four basic mathematical operations multiplication, subtraction, division, and addition – various functions which can be seen in Table 1 were used. Firstly, considering the effective parameter on sediment transport, six different models have been presented. Comparing the presented models with data sets which were not used in presenting models, the best model has been selected. To assess the accuracy of the models presented through GEP algorithm versus the existing equations, the experimental results of Ota and Nalluri [14] which had no role in the training of the GEP were used.

2. Non-deposition sediment transport equations

May et al. [17], with the use of seven different data sets [11,12,42–46] studied the existing sediment transport relations. Laboratory data was used to evaluate these relations. Results of studying the relations showed that each relation presents good results only for data sets which have been used for relation presenting, thus in order to present a relation for sediment transport studying at the limit of deposition, they presented following relation:

$$C_V = 3.03\times 10^{-2}(\frac{D^2}{A})(\frac{d}{D})^{0.6}(\frac{V^2}{g(s-1)D})^{1.5}(1-\frac{V_t}{V})^4 \tag{1}$$

$$V_t = 0.125[g(s-1)d]^{0.5}[\frac{y}{d}]^{0.47} \tag{2}$$

where C_V is volumetric sediment concentration, D pipe diameter, A Cross-sectional area of the flow, d median diameter of particle size, g gravitational acceleration, s specific gravity of sediment ($=\rho_s/\rho$), V flow velocity, V_t the required velocity for incipient motion of sediment (Eq. 2) and y flow depth.

In order to sediment transport at the limit of deposition Ackers et al. [18] considered the above relation as the best existing relation for designing usage and Vongvisessomjai et al. [15] too used Eq. 1 for verification of his relation. Considering volumetric sediment consideration (C_V) and relative flow depth (d/R), Ebtehaj et al. [47] presented the Fr in the form of following relations:

$$Fr = \frac{V}{\sqrt{g(s-1)d}} = 4.49 C_V^{0.21}\left(\frac{d}{R}\right)^{-0.54} \tag{3}$$

Ab. Ghani and Azamathulla [41] used GEP to predict the bed load transport in sewers. The authors presented their equation by considering the parameters of volumetric sediment concentration (C_V), the relative depth of flow (d/R), dimensionless particle number (D_{gr}) and Overall sediment friction factor ($\lambda_s=1.13D_{gr}^{0.01}C_V^{0.02}\lambda_C^{0.98}$, λ_C clear water friction factor) as follows:

$$\frac{V}{\sqrt{gd(s-1)}} = 1.425 + \left(\frac{-0.41}{(\frac{R}{d})} \right) + \left(\frac{\frac{C_V}{5.91} - 1}{D_{gr}} \right) + \left(\frac{0.014}{\lambda_s} + \lambda_s - 8.43 \lambda_s^{1.5} D_{gr} \frac{R}{d} \right) \tag{4}$$

3. Data collection

In this research, a combination of the lab test results by Vongvisessomjai et al. [15] and Ota and Nalluri [14] was used. The model is proposed using experimental results presented by Vongvisessomjai et al. [15], and the results of lab experiments are used to verify the feasibility of the model proposed by Ota and Nalluri [14]. Vongvisessomjai et al. [15] conducted their tests on pipes in two sizes of 100 and 150 mm in diameter and 16 m in length. They employed two sections to measure the flow: one at a distance of 4. 5 m upstream, and the other at the distance of 5.5 m downstream. These two points were 6 m apart. In each section the velocities were measured at flow surface, middle depth and near bottom and their mean average were taken as the average velocity. For the air/water phase of the flow, the Manning coefficient of roughness (*n*) was equal to 0.0125. Vongvisessomjai et al. [15] tests were conducted in a non-deposited bed state. More details are given in Vongvisessomjai et al. [15]. To validate the accuracy of results presented in this article, Ota and Nalluri [14] data were used for a limit of deposition. For the purpose of their tests at the limit of deposition, Ota and Nalluri [14] used six different dimensions of d (ranging from 0.71 mm to 5.61 mm). They conducted 24 tests in total. Moreover, to test the impact of granulation on sediment transport, they conducted 20 further experiments using five different ranges of sediments with an average diameter of d = 2 mm. More details are given in Ota and Nalluri [14]. Table 1 shows the range of the data used in their tests.

Table 1
Range of data in Ota and Nalluri [14] and Vongvisessomjai et al. [15] studies.

	y/D	*V* (m/s)	*R* (m)	C_V (ppm)	*d* (mm)
Ota and Nalluri (1999)	0.39-0.84	0.515-0736	0.005-0.076	16-59	0.6-6.3
Vongvisessomjai et al. (2010)	0.2-0.4	0.24-0.63	0.012-0.032	4 to 90	0.2-0.43

4. Overview of gene expression programming

Gene expression programming (GEP) is an expansion of genetic programming (GP) [48]. GEP belongs to the family of evolutionary algorithms and is closely related to genetic algorithms and genetic programming. From genetic algorithms it inherited the linear chromosomes of fixed length, and from genetic programming, it inherited the expressive parse trees of varied sizes and shapes [49]. The GEP procedure is such that initially required functions for model creation and terminal set are being selected. In the next step, in order to evaluate the aimed parameter (in this study *Fr*) and comparing it with the real value, existing data sets are being recalled. Afterward, in order to randomly present the initial population, chromosomes are being produced. In the next step, for population mass production with the use of existing chromosomes, the program is run, and the fitness of target function is surveyed. If we arrive at pause conditions, the program is stopped, otherwise with the use of new chromosomes - which have been corrected via genetic operators - as well as new population; again target function is being evaluated. This action continues until program pause conditions are present.

The fitness of an individual program (*i*) for fitness model (*j*) has been presented by Ferreira [50] in the following form:

$$If\ \ E(ij) \le p,\ then\ \ f_{(ij)} = 1,\ else\ \ f_{(ij)} = 0 \tag{5}$$

Where *p* precision and *E(ij)* the error of program *i* for fitness case (*j*). For the absolute error, it is being stated as in the following the form:

$$E(ij) = \left| p_{(ij)} - T_j \right| \tag{6}$$

Also, the fitness value (f_i) for an individual program is stated in the following form:

$$f_i = \sum (R - | p_{(ij)} - T_j) \tag{7}$$

where *R* is selection range, $p_{(ij)}$ the predicted value by the individual program (*i*) for fitness case (*j*) and T_j the target value for fitness case (*j*). After fitness function determination, the terminal set (*T*) and function set (*F*) have to be determined in order to select chromosomes.

5. Methodology

In order to survey sediment transport in pipes, effective parameters on flow and sediment particles movement have to be recognized. According to laboratory studies by researchers [12,15,17], the most important surveyed and utilized parameters to present their relations, include parameters like flow velocity (V), dimensionless particles number (D_{gr}), volumetric sediment concentration (C_V), median diameter of particles size (d), pipe diameter (D), flow depth (y), hydraulic radius (R), cross-sectional area of the flow (A), overall sediment friction factor (λ_s) and special gravity of sediment (s). Thus dimensionless parameters could be considered in the form of movement, transport, sediment, transport mode, and flow resistance. Movement parameters are respectively stated as densimetric Froude number (Fr) or (ψ) which uses shear stress instead of velocity. Transport parameter contains volumetric sediment concentration (C_V) or the presented transport parameter (φ), dimensionless particle number (D_{gr}), proportional average size of particles (d/D) and specific gravity of sediment (s). Transport form parameter includes the ratio of hydraulic radius to the median diameter of particles size (R/d), the ratio of squared pipe diameter to the flow cross-sectional area (D^2/A), relative flow depth (y/d) - instead of which usually R/d is being used - and the flow resistance parameter that considers flow overall frictional coefficient (λ_s). Based on these explanations, in order to study the effect of each and every parameter in different dimensionless groups, dimensionless parameters can be presented in order to predict Fr in the form of Table 2.

Table 2
Dimensionless sediment transport parameters in clean pipes.

Parameter type	Dimensionless groups
Movement	$Fr=\frac{V}{\sqrt{gd(s-1)}},\frac{1}{\psi}=\frac{\tau_0}{\rho g(s-1)d}$
Transport	$C_V,\varphi=\frac{C_V VR}{\sqrt{g(s-1)d^3}}$
Sediment	$D_{gr},d/D,s$
Transport mode	$d/R,D^2/A,d/y,y/D$
Flow resistance	$\lambda_s,(k_0-k_s)/D$

It is necessary to use different statistical indexes to verify the feasibility of the proposed model. The statistical indexes used in this study include dimensionless coefficient criteria called R-Squared (R^2), the relative criteria of Mean Average Percentage Error ($MAPE$) and absolute criteria of Root Mean Squared Error ($RMSE$). The R-Squared (R^2) index is the ratio of the combined dispersion of the estimated model and the observed value to the dispersion of the estimated and observed models. The $MAPE$ expresses the estimated value in relation to the observed value. $MAPE$ is a non-negative index which has no higher limit. The $RMSE$ is a criterion of mean error that has no upper limit and has the lowest possible value of zero, representing the best estimation by the model.

$$R^2 = \left[\frac{\sum_{i=1}^{n} \left(Fr_{EXP_i} - \overline{Fr_{EXP_i}} \right) \left(Fr_{GEP_i} - \overline{Fr_{GEP_i}} \right)}{\sqrt{\sum_{i=1}^{n} \left(Fr_{EXP_i} - \overline{Fr_{EXP_i}} \right)^2 \sum_{i=1}^{n} \left(Fr_{EXP_i} - \overline{Fr_{GEP_i}} \right)^2}} \right]^2 \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Fr_{EXP_i} - Fr_{GEP_i})^2} \tag{9}$$

$$MAPE = \left(\frac{100}{n}\right) \sum_{i=1}^{n} \left(\frac{| Fr_{EXP_i} - Fr_{GEP_i} |}{Fr_{EXP_i}} \right) \tag{10}$$

The indexes mentioned above present the estimated amounts as the average of the forecasted error and do not present any information on the forecasted error distribution of the suggested models. It is obvious that a high correlation coefficient (80- 90%) is not always considered as an indication of the high accuracy of a model; on the contrary, this index may lead to showing high accuracy for mediocre models [51]. In addition, *RMSE* index indicates the model's ability to predict a value away from the mean [52]. Therefore, the presented model must be evaluated using other indexes such as average absolute relative error (*AARE*) and threshold statistics [53–56]. TS_x index indicates forecasted error distribution by each model for x% of the anticipations. This parameter is determined for various amounts of average absolute relative error. The amount of the *TS* index for x% of the predictions is determined as explained below:

$$TS_x = \frac{Y_x}{n} \times 100 \tag{11}$$

$$AARE = \left(\frac{1}{n} \right) \sum_{i=1}^{n} \left(\frac{Fr_{EXP_i} - Fr_{GEP_i}}{Fr_{EXP_i}} \right) \tag{12}$$

Where Y_x is the number of the forecasted amounts of all the data for each amount of *AARE* is less than x%.

6. Derivation of densimetric Froude number based on GEP

This section concentrates on GEP method to calculate densimetric Froude number (*Fr*). The training set must be selected from amongst all the existing data to present a model. To that end, the data presented by Vongvisessomjai et al. [15] was selected as training set, and the data presented by Ota and Nalluri [14] was selected as a testing set. The training environment of the system has been defined after selecting the training set. After classifying the data, various parameters must be defined to make a model. To ,create the generation the initial population of the individuals, multi-genic chromosomes, are used which include four genes. We must now determine the number of the initial population. Considering Ferreira's [49] suggestion stating that using the size population within the range of 30- 100 can lead to good results, the size of the used population includes 50 chromosomes in this study which was selected through trial and error. After selecting the population size, the individuals are evaluated, and their fitness function is calculated using *MSE* as follows:

$$f_i = \frac{1000}{1+E_i} \quad for \quad E_i = P_{ij} - O_j \tag{13}$$

Where Q_{ij} is the amount observed for fitness case, and P_{ij} is the amount predicted by using i individual chromosome for fitness case j. The best state is when the equation E_{ij}= 0 is obtained. This means that the amounts predicted using i individual chromosome for fitness case j is equal to the amount observed for fitness case j (P_{ij}= E_{ij}). The set of terminals and the set of the function must be determined for each gene in the chromosome after selecting fitness function. The function sets used in this study include {×, -, ÷, ×, *Gau2*} while the set of terminals are as follows:

$$T = \left\{ F_r, C_V, D_{gr}, \frac{d}{D}, \frac{d}{R}, \frac{D^2}{A}, \frac{R}{D}, \lambda_s \right\} \tag{14}$$

Afterward, the number of genes and their head and tail length must be determined for each gene in the chromosome. By using trial and error and the succeeding rate, four genes were selected in the present study in each chromosome. The head length was selected to be 5 (*h*=5), and while the maximum number of arguments per function is equal to 2 (n_{max}= 2), the length of the tail turns out to be equal to 6 (*t*= 5 × (2-1) +1). The genetic operator rate must now be determined. Genetic operators such as mutation, inversion, transportation (IS, RIS, gene transportation), recombination and crossover (one point, two points, and gene recombination) were used. The rates of the mentioned parameters are presented in Table 3. We must finally determine the linking function. Considering the fact that using four different sub-expressions in this study has led to having four genes, the genes must be bound for us to reach the final result. Therefore, {+} operator has been used as the linking function among the genes in this study. Simulating the model begins after determining the essential parameters. *Gau2{x, y}* function presented in Table 2 returns $exp(-(x+y)^2)$ amount.

Table 3
Parameters of GEP model.

Parameter	Setting
Population size	50
Number of generations	250000
Number of chromosomes	50
Number of genes	4
function set	×, -, ÷, ×, Gau2
Linking function	Addition
Mutation rate	0.03
Inversion rate	0.15
IS transposition rate	0.1
RIS transposition rate	0.1
Gene transposition rate	0.15
One-point recombination rate	0.3
Two-point recombination rate	0.3
Gene recombination rate	0.15

7. Result and discussion

To study sediment transport, different parameters in no-deposition stage and to present a model which could estimate the best results in comparison with actual values, dimensionless parameters in Table 2 have been used. As considered in this table, dimensionless parameters that effect on the sediment transport in no-deposition mode are categorized into five groups. In order to present a model, the effect of four groups of transport, deposition, and transport form and flow resistance on movement group was surveyed. Thus, six different models are listed in Table 4. In the presented models, volumetric sediment concentration (C_V) which is related to transport dimensionless group and overall sediment frictional coefficient (λ_s) which is related to flow resistance dimensionless group have been considered constant. For sediment group D_{gr} and d/D parameters and transport form group, d/R ،D^2/A and y/D have been considered.

Table 4
Dependent parameters in predicting *Fr* considering the effect of dimensionless group parameters.

Model	Dependent parameter	Independent parameters	Train			Test		
			R^2	MAPE	RMSE	R^2	MAPE	RMSE
1	*Fr*	$C_V, D_{gr}, d/R, \lambda_s$	0.98	2.66	0.16	0.96	2.94	0.12
2	*Fr*	$C_V, D_{gr}, D^2/A, \lambda_s$	0.89	6.58	0.73	0.81	11.09	0.30
3	*Fr*	$C_V, D_{gr}, y/D, \lambda_s$	0.89	5.33	0.53	0.90	7.84	0.32
4	*Fr*	$C_V, d/D, d/R, \lambda_s$	**0.99**	**2.56**	**0.12**	**0.99**	**2.82**	**0.14**
5	*Fr*	$C_V, d/D, D^2/A, \lambda_s$	0.92	5.70	0.65	0.85	9.39	0.27
6	*Fr*	$C_V, d/D, y/D, \lambda_s$	0.97	2.94	0.20	0.96	3.05	0.19

Table 3 shows sextet presented models with the use of Table 1. In order to present models, laboratory results by Vongvisessomjai et al. [15] have been utilized. After presenting different models to evaluate the estimated results, via each model, the *Fr* has been surveyed with the use of Ota and Nalluri [14] laboratory results. According to verification criteria presented in Table 3, model 4 which uses volumetric sediment concentration (C_V), relative flow depth (*d/R*), proportional average size of particles (*d/D*) and overall frictional factor (λ_s) to estimate *Fr* delivers the best result. The *MAPE* index in *Fr* evaluation with the use of this model is about 2.56% for the test, and 2.82% for train and *RMSE* is 0.12 for train and 0.14 for the test. It is considered that the effect of data sets alternations on the precision of this model is about less than

1%. Other presented models in this table, compared to the mode with train data, show better results than test data, and this is an indication that using these models (models 1, 2, 3, 5 and 6) would not be trustworthy. Therefore, it could be said that to present a model which could well estimate *Fr* in a sewer at the limit of deposition state, effective parameters can be considered like model 4 in Table 3. This means that using C_V as transport parameter, *d/R* as transport form parameter, *d/D*as sediment parameter, and λ_s as flow resistance parameter in *Fr* evaluation, leads to good results. The presented equation through using the parameters of model 4 and expression tree presented in Figure 1 can be presented as follows. The amounts of the parameters presented in this figure have been shown in Table 5.

$$Fr = \left[\frac{d}{D} - \left(\left(89.66 - \left(\frac{d}{R}\right)\right) - \left(\frac{d}{D}\right)\right) + \frac{\lambda_s - \left(\frac{d}{R}\right)}{C_V + \left(\frac{d}{R}\right)}\right] + \left[\frac{33.1 \times C_V \times \left(\left(\frac{d}{D}\right) + 15.45\right)}{\frac{\lambda_s}{4.23} + \left(\frac{d}{R}\right) - \left(\frac{d}{D}\right)}\right] + \left[\frac{C_V}{\frac{\frac{d}{D} + 67 \times C_V}{\frac{\lambda_s}{\left(\frac{d}{D}\right)} + 132.42}}\right] + \left[exp\left(-\left(C_V - \lambda_s + \left(\frac{d}{D}\right)\right)^2\right) \times \left(\left(\frac{d}{D}\right) + 92.4\right) + \left(\left(\frac{d}{D}\right) + \frac{-5.9}{9.54}\right)\right] \tag{15}$$

We can rewrite the above-mentioned formula as follows:

$$Fr = \left[2 \times \left(\frac{d}{D}\right) + \left(\frac{d}{R}\right) - 89.66 + \frac{\lambda_s - \left(\frac{d}{R}\right)}{C_V + \left(\frac{d}{R}\right)}\right] + \left[\frac{33.1 \times C_V \times \left(\left(\frac{d}{D}\right) + 15.45\right)}{\frac{\lambda_s}{4.23} + \left(\frac{d}{R}\right) - \left(\frac{d}{D}\right)}\right] + \left[\frac{C_V \times \frac{\lambda_s}{\left(\frac{d}{D}\right)} + 132.42}{\left(\frac{d}{D}\right) + 67 \times C_V}\right] + \left[exp\left(-\left(C_V - \lambda_s + \left(\frac{d}{D}\right)\right)^2\right) \times \left(\left(\frac{d}{D}\right) + 92.4\right) + \left(\left(\frac{d}{D}\right) - 0.62\right)\right] \tag{16}$$

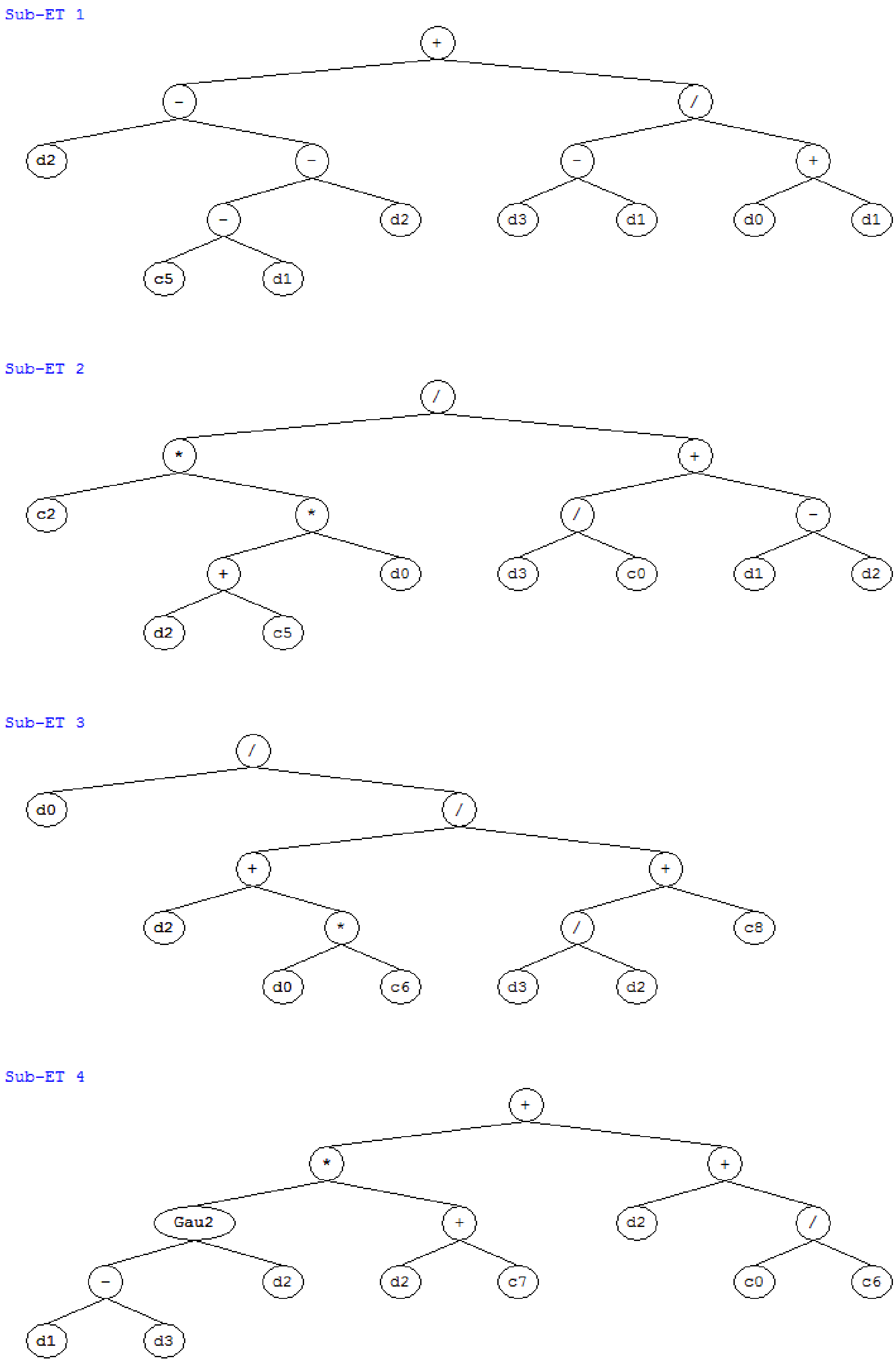

Fig. 1. Expression Tree (ET) for GEP formulation.

Table 5
Values of the parameters used in ET (Figure 1).

d0	d1	d2	d3	G1C5	G2C2	G2C0	G2C5	G3C8	G3C6	G4C7	G4C0	G4C6
C_v	d/R	d/D	λ_s	89.66	33.10	4.23	15.45	132.42	67.08	92.40	-5.90	9.54

Figure 2 shows the *Fr* results predicted by model 4 (Eq. 15) in both training and testing stage. Due to the fact that the accuracy of GEP model presented in Table 4 in this research (Eq. 15) has been studied quantitatively for both test (*MAPE*= 2.82 & *RMSE*= 0.14) and train (*MAPE*= 2.56 & *RMSE*= 0.12) states, in this figure, we will attend to studying the prediction results of the GEP model. The figure indicates that the forecasted *Fr* which were obtained through using GEP presented fairly good results in both train and test states while almost all estimated amounts have a relative error of less than 10%. The data used for the purpose of test and train of equation 15 have different ranges of *Fr* in such manner that the *Fr* used in training the model was within the range of 4 to 9 while the *Fr* range used in testing the model is 3 to 6. Therefore, it could be stated that while studying the model accuracy in test state all the *Fr* are not within the range which was used in training the model, thus, considering the qualitative results (Table 4) and quantitative results (Figure 2), it proves the accuracy of the presented results obtained by this equation.

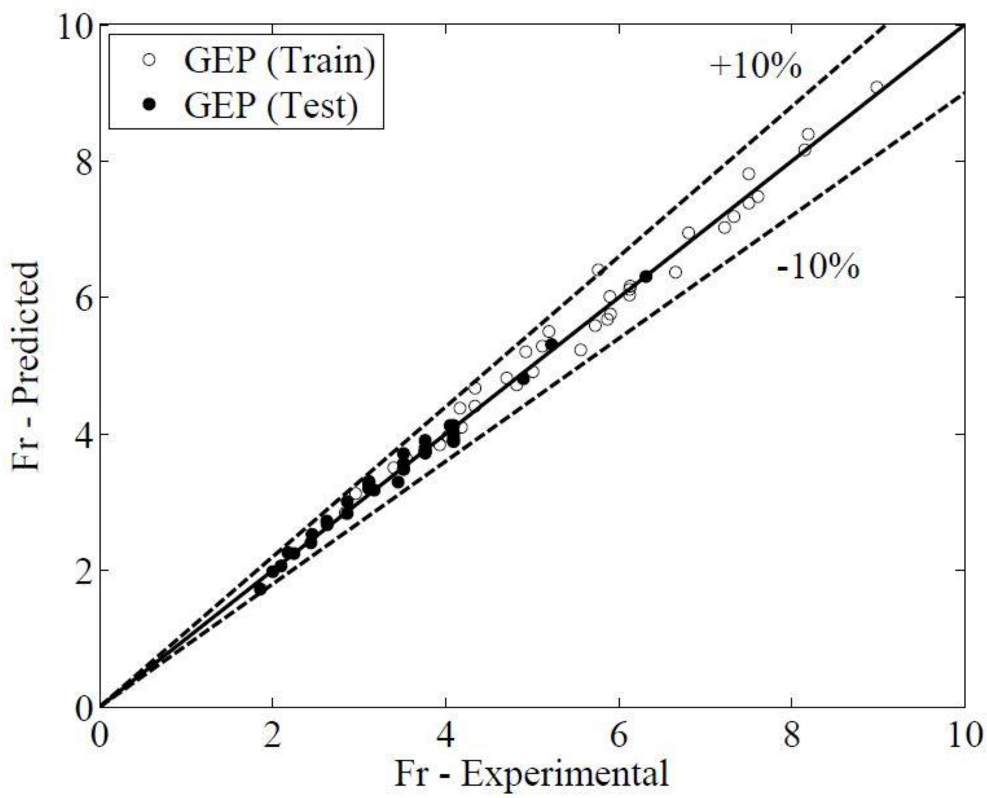

Fig. 2. Comparison of GEP result for both of train and test stages with actual values.

Figure 3 compares the prediction results of *Fr* by using the GEP model presented in this study (Eq. 15) and the existing regression equation with the actual values. The figure shows that the results of the presented predictions using GEP are almost accurate in a way that all forecasted points have an error less than ten percent and taking into consideration Figure 4 which shows the accumulative distribution of the error, we can see that the maximum relative error in estimating *Fr* through using GEP is almost equal to 6 %. Also, this figure indicates that approximately 90% of the anticipated amounts have a relative error of less than 5%. Now in case, we intend to study the results of the presented model through statistical indexes, referring to Table 6 shows us that the amounts of the presented statistical indexes for this model with an R^2= 0.99, *MAPE*= 2.82 and *RMSE*= 0.14 is minimum in an amount in comparison to other equations presented in this table. The equation presented by Ab. Ghani and Azamathulla [41] is less accurate (R^2= 0.74, *MAPE*= 13.18 and *RMSE*= 0.49) considering Table 6 and Figure 4. The figure indicates that in the majority of the points the results are presented with an error more than 10 percent.

Figure 4 shows that only 25% of the amounts estimated by this model have a relative error of less than 10%. Also, it indicates that some of the *Fr* forecasted by this equation have a relative error of more than 30 percent which indicates the uncertainty of the equation presented by Ab Ghani and Azamathulla [41]. Therefore, using this equation for the purpose of estimating *Fr* cannot be that much confidence. Ebtehaj et al. [47] equation are fairly accurate because it estimates the majority of *Fr* with a less-than-10-percent relative error, but it is less accurate in comparison with the equation presented in this study. This is in a way that considering Figure 4, which indicates the distribution of the estimation error by different models, we can see that approximately 70 percent of the estimation results of this model have an error less than 5% while for the model presented in this study the predicted amounts have an error of less than 5% for almost 90% of the *Fr*. At times, May et al. [17] equation which has been obtained through semi-experimental method and has been known as one of the best sediment transport equations in limit of deposition [15] presents the estimated amounts with a more- than-15% relative error according to Figure 5 while the equation presented in this study has a maximum relative error of 6%. Also, considering Figure 6, the amounts of statistical indexes presented by this equation (R^2= 0.93, *MAPE*= 5.74 and *RMSE*= 0.24) indicates lesser estimation accuracy of this equation in comparison with that of the presented equation.

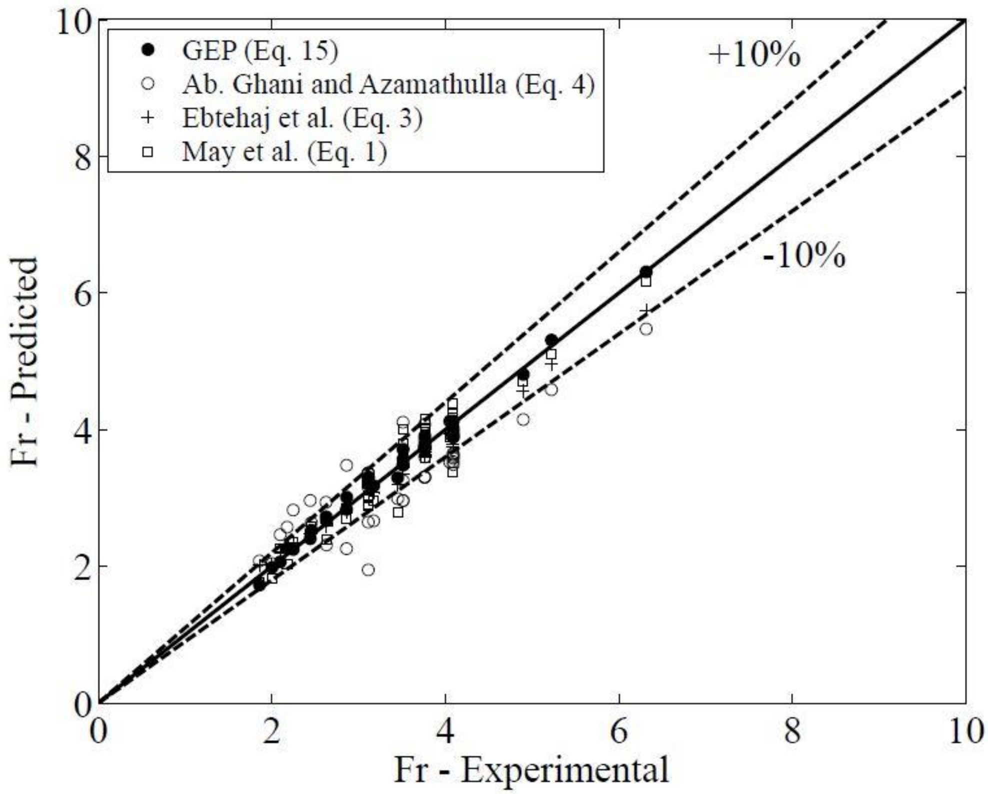

Fig. 3. Comparison of proposed equation and existing equations.

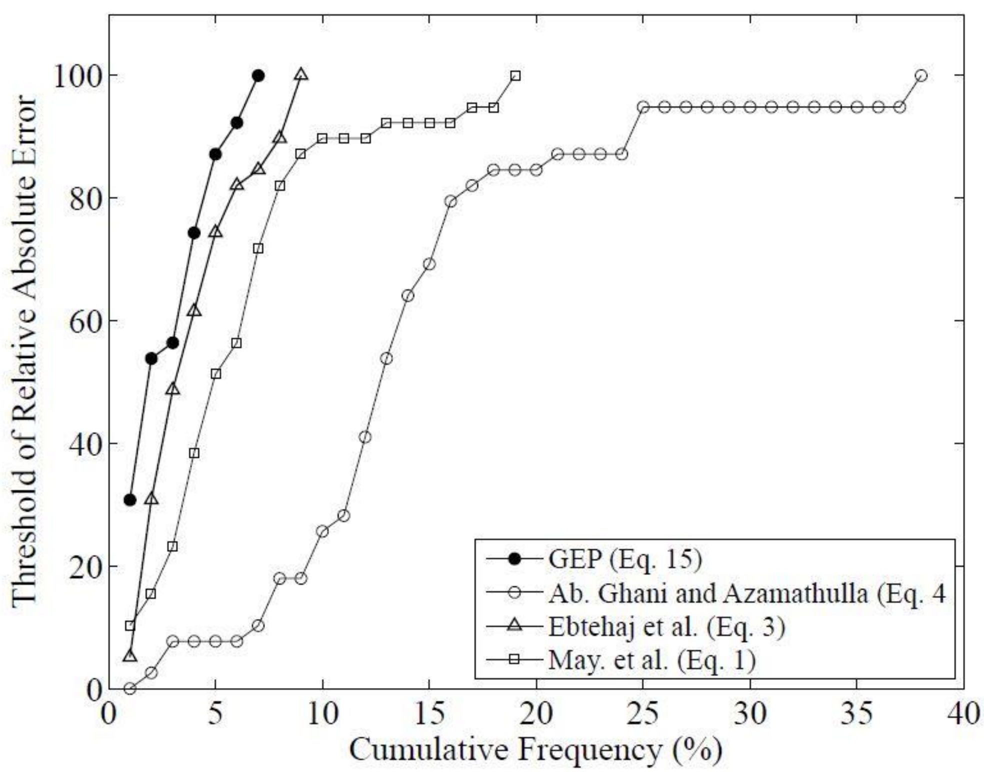

Fig. 4. Error distribution of for GEP and existing equations.

Table 6
Validation of proposed equation and existing equations with statistical indexes.

Equation	R^2	*MAPE*	*RMSE*
Proposed Equation (Eq. 15)	**0.99**	**2.82**	**0.14**
Ab. Ghani and Azamathulla (Eq. 4)	0.74	13.18	0.49
Ebtehaj et al. (Eq. 3)	0.97	3.70	0.18
May et al. (Eq. 1)	0.93	5.74	0.24

Accordingly, in this study, the effects of GEP model output on the variations of dimensionless particle number (D_{gr} in this study) were investigated. The discrepancy ratio (*DR*) (ratio of predicted to actual values) was employed to measure the sensitivity of the GEP model to D_{gr} parameter. A *DR* value of 1 shows a perfect agreement, while values smaller (or greater) than 1 indicate under (or over) prediction of discharge coefficient inside weir. The result of the GEP model for variations of the discrepancy ratio (*DR*) values is plotted versus the dimensionless particle number (D_{gr}) in Figure 5. The maximum, mean and minimum *DR* values for GEP model were obtained 1.06, 1.005 and 0.93, respectively. As Figure 5 shows it, for almost all the *Fr* estimated, the *DR* is close to 1. When GEP predicts the model using the over design way, the dimensionless particle number will be equal to 1.7 (*DR*= 1.06), and when it uses the under-design way to predict *Fr*, dimensionless particle number is equal to 2.15 (*DR*= 0.93).

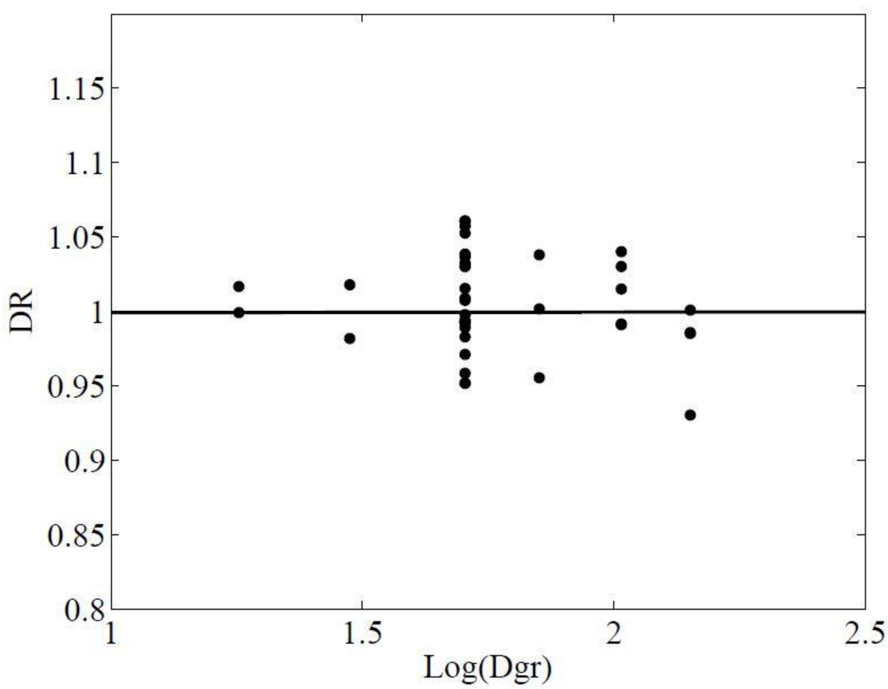

Fig. 5. DR values versus D_{gr} for GEP model.

Based on explanations given in Figures 3 and 5, and Table 6, the equation presented in this study is more accurate than the existing regression equations. While it is more accurate in studying the estimation accuracy through using statistical indexes, it is also more accurate in studying the estimation error distribution in Figure 4 and can be utilized as a substituting method in *Fr* estimation for sediment transport in no-deposition mode.

8. Conclusion

Transmitting flow from sewerage systems often contains suspended materials. Therefore, transporting suspended materials and preventing their sedimentation are important matters inflow transport through sewerage networks. Different methods have been presented for sediment transport in sewage, but due to the lack of recognition of effective factors on sediment transport, these methods show different results in different conditions. Hence, in recent years, soft computations have been utilized in order to estimate densimetric Froude number (*Fr*) in these systems. In this paper, with the use of the presented model by Gene-expression programming (GEP), *Fr* has been estimated. In order to present the effective factor on *Fr* estimation, six different models were presented. In these models, the effect of movement, transport, sediment, transport mode and flow resistance parameters have been considered. After *Fr* estimation, the precision of all sextet models has been studied. The results indicated that among the three parameters provided by "Transport mode" group, the best and the worst accuracy were achieved by using d/R and D^2/A (respectively) as improper use of the parameters of this group, up to two-fold increase relative error. In addition to, Also, with the constant parameters in the groups "transport", "flow resistance" and "transport mode", the parameter d/D in all input combinations, leading to better results than when used D_{gr} as "sediment" parameter. Therefore, it was revealed that the model which considers volumetric sediment concentration (C_V), relative flow depth (d/R), proportional average size of particles (d/D), overall friction factor (λ_s) for *Fr* estimation, shows the best results. The presented model estimates *Fr* with an average error value of about 2.82%. The comparison of existing methods illustrated the high level of accuracy of Ebtehaj *et al.* (Eq. 3) method in comparison with others. It should not be an inappropriate use of GEP functions such as Eq. (4) results in weak performance of the model. The presented model with existing values was also studied, and the results showed that in proportion with existing relations the model well estimates the *Fr*. Incidentally making use of the proposed GEP-based technique in the form of the most superlative formulations has a dominant role to experience in the attaining astonishing and remarkable successes for real-world application. Another plus aspect of this study is the use of extracted mathematical expressions as a realistically valuable technique for practical engineering as an alternative for existing methods.

Notation

A	Cross-sectional area of the flow
C_V	Volumetric sediment concentration
D	Pipe diameter
d	Median diameter of particle size
D_{gr}	Dimensionless particle number
$E_{(ij)}$	Error of program i for fitness case j (Eq. 5)
Fr	Densimetric Froude number
P	Precision (Eq. 5)
$P_{(ij)}$	Value predicted by individual program i for fitness case j (Eq. 6)
R	Hydraulic radius, Selection Range (Eq. 6)
s	Specific gravity of sediment ($=\rho_s/\rho$)
V	Velocity of flow
V_t	Incipient flow velocity which follows from equation (2)
y	Flow depth
λ_c	Clear Water friction factor
λ_s	Overall sediment friction factor
ψ	Flow parameter
φ	Transport parameter

References

[1] Butler D, Clark P. Sediment management in urban drainage catchments. Report No. 134, Construction Industry Research and Information Association, London, UK: 1995.

[2] ASCE. Water pollution control federation: Design and construction of sanitary and storm sewers. American Society of Civil Engineers Manuals and Reports on Engineering Practices, No. 37, Reston, VA: 1970.

[3] BS8005-1. Sewerage Guide to New Sewerage Construction. British Standard Institution, London, UK: 1987.

[4] European Standard EN 752-4. Drain and sewer system outside building: Part 4. Hydraulic design and environmental considerations. Brussels: CEN (European Committee for Standardization): Part; 1997.

[5] Lysne DK. Hydraulic design of self-cleaning sewage tunnels. J Sanit Eng Div 1969;95:17–36.

[6] Pedroli R. Bed load transportation in channels with fixed and smooth inverts. Scuola Politecnica Federale, Zurigo, Switzerland, 1963.

[7] GRAF WH, ACAROGLU ER. Sediment transport in conveyance systems (Part 1)/A physical model for sediment transport in conveyance systems. Hydrol Sci J 1968;13:20–39.

[8] Novak P, Nalluri C. Sediment transport in smooth fixed bed channels. J Hydraul Div 1975;101:1139–54.

[9] Nalluri C. Sediment transport in rigid boundary channels. Proceeding Euromech 192 Transp. Suspended Solids Open channels, vol. 192, Neubiberg, Germany: 1985, p. 101–4.

[10] Mayerle R. Sediment transport in rigid boundary channels. University of Newcastle Upon Tyne, UK, 1988.

[11] Mayerle R, Nalluri C, Novak P. Sediment transport in rigid bed conveyances. J Hydraul Res 1991;29:475–95. doi:10.1080/00221689109498969.

[12] Ghani A. Sediment transport in sewers. University of Newcastle Upon Tyne, UK, 1993.

[13] Azamathulla HM, Ab. Ghani A, Fei SY. ANFIS-based approach for predicting sediment transport in clean sewer. Appl Soft Comput 2012;12:1227–30. doi:10.1016/j.asoc.2011.12.003.

[14] Ota JJ, Nalluri C. Graded sediment transport at limit deposition in clean pipe channel. 28th Congr. Int. Assoc. Hydro-Environmental Eng. Res. Graz, Austria, 1999.

[15] Vongvisessomjai N, Tingsanchali T, Babel MS. Non-deposition design criteria for sewers with part-full flow. Urban Water J 2010;7:61–77. doi:10.1080/15730620903242824.

[16] May RWP. Sediment transport in sewers. Hydraulic Research Station, Wallingford, England, Report IT 222; 1982.

[17] May RWP, Ackers JC, Butler D, John S. Development of design methodology for self-cleansing sewers. Water Sci Technol 1996;33:195–205.

[18] Ackers JC, Butler D, May RWP. Design of sewers to control sediment problems. Report No. CIRIA 141, Construction Industry Research and Information Association, London, UK: Construction Industry Research and Information Association London; 1996.

[19] Ackers P, White WR. Sediment transport: new approach and analysis. J Hydraul Div 1973;99:2041–60.

[20] Ackers P. Sediment aspects of drainage and outfall design. Proc. Int. Symp. Environ. Hydraul. Hong kong, 1991.

[21] Butler D, May R, Ackers J. Self-Cleansing Sewer Design Based on Sediment Transport Principles. J Hydraul Eng 2003;129:276–82. doi:10.1061/(ASCE)0733-9429(2003)129:4(276).

[22] Banasiak R. Hydraulic performance of sewer pipes with deposited sediments. Water Sci Technol 2008;57:1743. doi:10.2166/wst.2008.287.

[23] Ota JJ, Perrusquía GS. Particle velocity and sediment transport at the limit of deposition in sewers. Water Sci Technol 2013;67:959. doi:10.2166/wst.2013.646.

[24] Safari M-J-S, Aksoy H, Unal NE, Mohammadi M. Non-deposition self-cleansing design criteria for drainage systems. J Hydro-Environment Res 2017;14:76–84. doi:10.1016/j.jher.2016.11.002.

[25] Mondal SK, Jana S, Majumder M, Roy D. A comparative study for prediction of direct runoff for a river basin using geomorphological approach and artificial neural networks. Appl Water Sci 2012;2:1–13. doi:10.1007/s13201-011-0020-3.

[26] GAD MI, Khalaf S. Application of sharing genetic algorithm for optimization of groundwater management problems in Wadi El-Farigh, Egypt. Appl Water Sci 2013;3:701–16. doi:10.1007/s13201-013-0114-1.

[27] Al-Abadi AM. Modeling of stage–discharge relationship for Gharraf River, southern Iraq using backpropagation artificial neural networks, M5 decision trees, and Takagi–Sugeno inference system technique: a comparative study. Appl Water Sci 2016;6:407–20. doi:10.1007/s13201-014-0258-7.

[28] Ahmadianfar I, Adib A, Taghian M. Optimization of multi-reservoir operation with a new hedging rule: application of fuzzy set theory and NSGA-II. Appl Water Sci 2017;7:3075–86. doi:10.1007/s13201-016-0434-z.

[29] Azimi H, Bonakdari H, Ebtehaj I, Ashraf Talesh SH, Michelson DG, Jamali A. Evolutionary Pareto optimization of an ANFIS network for modeling scour at pile groups in clear water condition. Fuzzy Sets Syst 2017;319:50–69. doi:10.1016/j.fss.2016.10.010.

[30] Ebtehaj I, Bonakdari H. Evaluation of Sediment Transport in Sewer using Artificial Neural Network. Eng Appl Comput Fluid Mech 2013;7:382–92. doi:10.1080/19942060.2013.11015479.

[31] Ebtehaj I, Bonakdari H. Performance Evaluation of Adaptive Neural Fuzzy Inference System for Sediment Transport in Sewers. Water Resour Manag 2014;28:4765–79. doi:10.1007/s11269-014-0774-0.

[32] Rezaei H, Rahmati M, Modarress H. Application of ANFIS and MLR models for prediction of methane adsorption on X and Y faujasite zeolites: effect of cations substitution. Neural Comput Appl 2017;28:301–12. doi:10.1007/s00521-015-2057-y.

[33] Singh R, Vishal V, Singh TN. Soft computing method for assessment of compressional wave velocity. Sci Iran 2012;19:1018–24. doi:10.1016/j.scient.2012.06.010.

[34] Singh R, Vishal V, Singh TN, Ranjith PG. A comparative study of generalized regression neural network approach and adaptive neuro-fuzzy inference systems for prediction of

unconfined compressive strength of rocks. Neural Comput Appl 2013;23:499–506. doi:10.1007/s00521-012-0944-z.

[35] Isanta Navarro R. Study of a neural network-based system for stability augmentation of an airplane. Univ Politècnica Catalunya 2013.

[36] Gan Z, Yang Z, Li G, Jiang M. Automatic Modeling of Complex Functions with Clonal Selection-Based Gene Expression Programming. Third Int. Conf. Nat. Comput. (ICNC 2007), IEEE; 2007, p. 228–32. doi:10.1109/ICNC.2007.278.

[37] Azamathulla HM, Ghani AA. Genetic Programming to Predict River Pipeline Scour. J Pipeline Syst Eng Pract 2010;1:127–32. doi:10.1061/(ASCE)PS.1949-1204.0000060.

[38] Khan M, Azamathulla HM, Tufail M, Ab Ghani A. Bridge pier scour prediction by gene expression programming. Proc Inst Civ Eng - Water Manag 2012;165:481–93. doi:10.1680/wama.11.00008.

[39] CHANG CK, AZAMATHULLA HM, ZAKARIA NA, GHANI AA. Appraisal of soft computing techniques in prediction of total bed material load in tropical rivers. J Earth Syst Sci 2012;121:125–33. doi:10.1007/s12040-012-0138-1.

[40] Azamathulla HM, Ahmad Z. Gene-expression programming for transverse mixing coefficient. J Hydrol 2012;434–435:142–8. doi:10.1016/j.jhydrol.2012.02.018.

[41] Ab. Ghani A, Md. Azamathulla H. Gene-Expression Programming for Sediment Transport in Sewer Pipe Systems. J Pipeline Syst Eng Pract 2011;2:102–6. doi:10.1061/(ASCE)PS.1949-1204.0000076.

[42] Macke E. About sedimentation at low concentrations in partly filled pipes. (In Ger Mitteilungen, LeichtweissÐInstitut FuÈ r Wasserbau Der Tech Univ t Braunschweig 1982.

[43] May RWP, Brown PM, Hare GR, Jones KD. Self-cleaning conditions for sewers carryign sediment. Hydraulic Research Ltd (Wallingford), Report SR 221; 1989.

[44] May RWP. Sediment transport in pipes, sewers and deposited beds. Hydraulic Research Ltd., Wallingford, England, Report SR 320; 1993.

[45] Nalluri C, Ab-Ghani A. Bed load transport with deposition in channels of circular cross section. Proc. Sixth Int. Conf. Urban Storm Drain., Niagara Falls, Canada: 1993.

[46] Nalluri C, Ghani AA, El-Zaemey AKS. Sediment transport over deposited beds in sewers. Water Sci Technol 1994;29:125–33.

[47] Ebtehaj I, Bonakdari H, Sharifi A. Design criteria for sediment transport in sewers based on self-cleansing concept. J Zhejiang Univ Sci A 2014;15:914–24. doi:10.1631/jzus.A1300135.

[48] John R. Koza. Genetic programming: on the programming of computers by means of natural selection. MIT Press Cambridge, MA, USA; 1992.

[49] Ferreira C. Algorithm for solving gene expression programming: a new adaptive problems. Complex Syst 2001;13:87–129.

[50] Ferreira C. Gene expression programming: mathematical modeling by an artificial intelligence. vol. 21. 2nd ed. Springer; 2006.

[51] Legates DR, McCabe GJ. Evaluating the use of “goodness-of-fit” Measures in hydrologic and hydroclimatic model validation. Water Resour Res 1999;35:233–41. doi:10.1029/1998WR900018.

[52] Hsu K, Gupta HV, Sorooshian S. Artificial Neural Network Modeling of the Rainfall-Runoff Process. Water Resour Res 1995;31:2517–30. doi:10.1029/95WR01955.

[53] Jain A, Kumar Varshney A, Chandra Joshi U. Short-Term Water Demand Forecast Modelling at IIT Kanpur Using Artificial Neural Networks. Water Resour Manag 2001;15:299–321. doi:10.1023/A:1014415503476.

[54] Jain A, Ormsbee LE. Short-term water demand forecast modeling techniques-CONVENTIONAL METHODS VERSUS AI. J Am Water Works Assoc 2002;94:64–72. doi:10.1002/j.1551-8833.2002.tb09507.x.

[55] Rajurkar MP, Kothyari UC, Chaube UC. Modeling of the daily rainfall-runoff relationship with artificial neural network. J Hydrol 2004;285:96–113. doi:10.1016/j.jhydrol.2003.08.011.

[56] Maghrebi MF, Givehchi M. Using non-dimensional velocity curves for estimation of longitudinal dispersion coefficient. Proc. seventh Int. Symp. river Eng., 2007, p. 16–8.